

## REMARKS

Claims 1, 10, and 19 have been amended to clarify the subject matter regarded as the invention. Claims 1-6 and 8-27 remain pending. (Applicant notes that claim 7 was canceled in Amendment A.)

The Examiner has rejected claims 1-6 and 7-27 under 35 U.S.C. §103(a) as being unpatentable over Sitrick in view of Ginter.

The rejection is respectfully traversed. With respect to claim 1, the claim recites “extracting the person image portion of the received video image,” “recognizing an identity of the user based on said person image of the user by matching the person image of the user with an image stored in a user image database,” and “selecting a subset of the vision-enabled content based on the identity of the user as recognized by matching the person image of the user with an image stored in a user image database.” The present application supports the above-quoted limitations by describing, without limitation, selecting the content to be sent to a user based on such image-based recognition (Application at 10:11-27) and matching a user with other users of similar skill based on such image-based recognition of a user (Application at 15:19-24).

Sitrick teaches, by contrast, determining the content to be provided to a user either by reading a storage medium inserted by a user, such as a game cartridge, or by receiving a user input, Sitrick at 23:25-35, not by “recognizing an identity of the user” by comparing an extracted “person image” with one stored in a database and determining which content to be provided based on the identity of the user as recognized by such image processing techniques. The Office Action notes that Sitrick teaches storing “user visual image data” in memory, Sitrick at 26:53-60, but Sitrick describes storing that data for purposes of integrating portions of such data into a

video game, see, e.g., Sitrick at 25:36-42, not later recognizing the identity of a user based on image processing as recited in claim 1.

Ginter is relied on in the Office Action solely for financial elements of claim 1 and does not appear to describe “extracting the person image portion of the received video image,” “recognizing an identity of the user based on said person image of the user by matching the person image of the user with an image stored in a user image database,” and “selecting a subset of the vision-enabled content based on the identity of the user as recognized by matching the person image of the user with an image stored in a user image database,” as recited by claim 1. As such, claim 1 is believed to be allowable.

Claims 2-6 and 8-9 depend from claim 1 and are believed to be allowable for the same reasons described above.

Claim 10 recites providing a program that “extracts from each video image the associated person image of the user to create a series of person images,” “processes the series of person images to detect a movement by said user,” and “controls the vision-enabled content based on said movement.” The word “action” has been amended to read “movement” to further clarify that the series of person images are processed to detect a physical movement by the user, such as a gesture or other physical movement, and the detected movement used to control the vision-enabled content. See, e.g., Application at 13:27-14:7 (movements as detected by image processing used to scroll up or down) & 16:24-26 (movement as detected used to control movements of a character in a video game). To control content, Sitrick teaches using traditional input devices, such as a mouse, joystick, keyboard or other device; using well-known virtual reality equipment, such as helmets, goggles, gloves, and other motion detectors worn on the body and configured to generate signals corresponding to movements of the detector; or using

biometric devices attached to the body and configured to monitor one or more physical parameters. Sitrick 21:3-23; 27:35-52; 34:40-50; 35:23-31. Sitrick describes storing and processing user image data, but solely for purposes of integrating such image data into a video game, see, e.g., Sitrick at 13:34-48; not for controlling video content by using image processing to detect user movements. As such, claim 10 is believed to be allowable over Sitrick.

Likewise, Ginter does not appear to teach providing a program that “extracts from each video image the associated person image of the user to create a series of person images,” “processes the series of person images to detect a movement by said user,” and “controls the vision-enabled content based on said movement,” as recited in claim 10. Therefore, claim 10 is believed to be allowable.

Claims 11-18 depend from claim 10 and are believed to be allowable for the same reasons described above.

Like claim 1, claim 19 recites, “receiving a video image comprising a person image of a user,” “recognizing an identity of the user based on said person image of the user by matching the person image of the user with an image stored in a user image database,” and “selecting a subset of the vision-enabled content based on the identity of the user as recognized by matching the person image of the user with an image stored in a user image database.” As such, claim 19 is believed to be allowable for the same reasons described above with respect to claim 1.

Claims 20-25 depend from claim 19 and are believed to be allowable for the same reasons described above.

Similarly to claim 10, claim 26 recites, “receiving a series of images of the user,” “recognizing a person image of the user in at least two images comprising the series of images,”

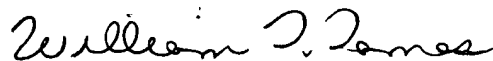
and “controlling the content based on the person image by detecting an action by the user based on changes in the person image between the at least two images.” As such, claim 26 is believed to be allowable for the same reasons described above with respect to claim 10. In particular, to the extent that Sitrick describes detecting an action by a user, Sitrick teaches doing so using virtual reality technology, such as gloves, helmets, and other devices worn on the person of the user that comprise motion detector devices, not by “detecting an action by the user based on changes in the person image between the at least two images,” as recited in claim 26.

Claim 27 depends from claim 26 and is believed to be allowable for the same reasons described above.

Attached hereto is a marked-up version of the changes made to the specification and claims by the current amendment with additions underlined and deletions struck through. The attached page is captioned “Version with markings to show changes made.”

Reconsideration of the application and allowance of all claims are respectfully requested based on the preceding remarks. If at any time the Examiner believes that an interview would be helpful, please contact the undersigned.

Respectfully submitted;



William J. James  
Registration No. 40,661  
V 408-973-2592  
F 408-973-2595

VAN PELT AND YI, LLP  
10050 N. Foothill Blvd., Suite 200  
Cupertino, CA 95014

**VERSION WITH MARKINGS TO SHOW CHANGES MADE**

**AMENDMENTS TO THE CLAIMS**

1. (Amended Three Times) A method of conducting commerce over a network, comprising:
  - encoding content for conversion into vision-enabled content;
  - receiving payment for encoding the content;
  - providing a program to decode the vision-enabled content;
  - receiving a video image comprising a person image of a user;
  - extracting the person image portion of the received video image;
  - recognizing an identity of the user based on said person image of the user by matching the person image of the user with an image stored in a user image database;
  - selecting a subset of the vision-enabled content based on the identity of the user as recognized by matching the person image of the user with an image stored in a user image database; and
  - sending the selected subset of the vision-enabled content to the user over a network, wherein the program decodes the selected subset of the vision-enabled content and combines the image of the user with the selected subset of the vision-enabled content.
  
10. (Amended Three Times) A method of conducting commerce over a network, comprising:
  - encoding content for conversion into vision-enabled content;
  - receiving payment for encoding the content;
  - providing a program to decode the vision-enabled content; and
  - sending the vision-enabled content to a user over a network, wherein the program:
    - decodes the vision-enabled content;
    - receives a series of video images, each image comprising a person image of the user;
    - extracts from each video image the associated person image of the user to create a series of person images; and

processes the series of person images to detect [an action] a movement by said user; and  
controls the vision-enabled content based on said [action] movement.

19. (Amended Three Times) A method of conducting commerce over a network, comprising:  
encoding content for conversion into vision-enabled content;  
providing a program to decode the vision-enabled content;  
receiving a video image comprising a person image of a user;  
recognizing an identity of the user based on said person image of the user by  
matching the person image of the user with an image stored in a user image database;  
selecting a subset of the vision-enabled content based on the identity of the user as  
recognized by matching the person image of the user with an image stored in a user  
image database; and  
sending the selected subset of the vision-enabled content to the user over a  
network, wherein the program decodes the selected subset of the vision-enabled content.